The Manchester School 114–134 Supplement 2010
doi: 10.1111/j.1467-9957.2010.02204.x

INVESTIGATION OF STOCHASTIC PAIRS TRADING STRATEGIES UNDER DIFFERENT VOLATILITY REGIMES

by
SAYAT R. BARONYAN
Informatics Institute, Computational Science & Engineering Graduate Program, Istanbul Technical University
İ. İLKAY BODUROĞLU
Namik Kemal University
and
*EMRAH ŞENER**
Center for Computational Finance, Ozyeğin University

We investigate several market-neutral trading strategies and find empirical evidence that market-neutral equity trading outperforms in 2008, the first full year of the global financial meltdown. In our experiments we use 14 distinct market-neutral trading strategies, using the combination of seven trading methods and two selection methods of pairs trading.

1 Introduction

Market-neutral equity trading strategies exploit mispricings in a pair of similar stocks (Beliossi, 2002). Mispricing is more usual in a global financial crisis (Gatev *et al.*, 2006). Therefore, more possibilities emerge at bad times. Moreover, there are fewer market participants, which reduces competition. Therefore, it is not surprising that market-neutral trading overperforms during most severe market conditions. In this paper, we propose several market-neutral equity trading strategies and find empirical evidence for the above statement using several trading strategies.

We propose several market-neutral equity trading systems (also known as pairs trading) and show that not only do they outperform existing systems, but they also beat the global financial crisis of 2008 by bringing in a more than 40 per cent net annual profit. Each and every one of selected market-neutral equity trading systems uses a combination of the following well-known tools of econometrics to select market-neutral pairs: augmented Dickey–Fuller (ADF) (Dickey and Wayne, 1979) and Granger causality tests (Granger, 1969) along with beta calculation in order to select market-neutral pairs. Furthermore, we use the Vasicek stochastic differential equation (Vasicek, 1977) for modeling the dynamics of the ratio of prices of a pair of

* The authors are indebted to Wolfgang Hörmann and Refik Gullu for suggesting ideas to better present the material. We also thank Sevda Akyüz who have proofread the paper. Finally, we are grateful to the three anonymous referees along with the Guest Editors Turalay Kenc and George Bratsiotis for their valuable suggestions, which helped us to understand the nature of the problem much better.

Published by Blackwell Publishing Ltd, 9600 Garsington Road, Oxford OX4 2DQ, UK, and 350 Main Street, Malden, MA 02148, USA.

stocks, and generalized method of moments (GMM) (Hansen, 1982), a non-parametric method, for parameter estimation of the Vasicek method. On top of these, we investigate a very simple trading strategy that proves to be low-risk and high-return seen from the perspective of our disjoint training and backtesting results that span 10 years ending on the last trading of December 2008.

The seminal pairs trading implementation was developed in the late 1980s by quantitative analysts led by Gerald Bamberger at Morgan Stanley. Pairs trading is a well-known trading idea that involves taking one long and one short position in two assets A and B whose prices P_A and P_B are believed to have a ratio $R_t = P_A/P_B$ that is mean-reverting over time. If, for instance, the spread $P_A – P_B$ is much greater than usual, then one would expect it to diminish in such a way as to make the ratio return to its long-term average, θ. Taking a long position in asset B along with a short position, with an equal dollar amount, in asset A constitutes a pair-trade. Independent of the direction of P_A, P_B or the market, once R_t returns to θ, the trade is closed with a profit.

One of the most notable papers on pairs trading is written by Gatev *et al.* (2006), who offer a comprehensive analysis. The authors use daily US data from 1962 to 2002 and show that a simple pairs trading rule produces excess returns of 11 per cent per annum and a monthly Sharpe ratio which can be up to six times larger than market returns. It is shown that the returns have high risk adjusted alphas, low exposure to known sources of systematic risk, cover reasonable transaction costs, and do not come from contrarian relative-price momentum strategies as documented in Lehmann (1990). However, the returns are comparable in magnitude to relative-price momentum strategies explained in Jegadeesh and Titman (1993). Gatev *et al.* interpret pairs trading profits as pointing towards a systematic dormant factor relating to the agency costs of professional arbitrage. The minimum distance criterion, based on the 'law of one price', is proposed as a metric to select the best pairs. It is argued that the important economic principle of the law of one price explains these arbitrage profits. The information period length is determined to be a constant 60 days.

In addition, Kovajecz and Odders-White (2004) link such high returns with market making trading activities, which allow price discovery of the underlying securities. Avellaneda and Lee (2008), on the other hand, focus on investigating two different trading signals in constructing principal component analysis based and exchange traded fund based strategies. Also, regarding the trading volumes of stocks, the paper analyzes the performance of these statistical arbitrage strategies. As in Gatev *et al.* the information period length is determined to be a constant 60 days. Mostly focusing on the selection criteria of pairs, Vidyamurthy (2004) and Herlemont (2003) propose the use of cointegration (Engle and Granger, 1987). Elliott *et al.* (2005) explicitly model the mean reversion process of the difference between the prices of paired stocks in continuous time. Also Perlin (2007) reported that pairs trading was a profitable strategy at the Brazilian market.

Of course, the risk that one takes when entering a pair-trade is the possibility of a structural breakdown of the mean-reverting-ratio property. This paper tries to minimize this risk by choosing the pairs more carefully. We describe efficient methods for selecting a number of pairs and then discuss rules for entering and exiting pair-trades.

A well-constructed pairs trading system needs to possess the following four basic properties:

- a reliable pair selection criterion;
- an efficient stochastic model to mimic the motion of a given pair;
- a good parameter estimation technique for the stochastic model;
- a low-risk high-return trading strategy with the given pair(s);
- a comprehensive backtesting methodology using disjoint training and testing data going back at least 10 years.

Our methodology includes the following properties.

- We propose a number of new market-neutral pair selection rules where we pick the best five pairs that pass different combinations of the ADF test, two-way Granger causality test and the market factor ratio (MFR) test, which we define later in the paper.
- We use GMM, a non-parametric method, for parameter estimation of the Vasicek model in market-neutral trading for the first time. Note that the Vasicek model has been used in pairs trading earlier in an unpublished paper (Do *et al.*, 2006). However, its parameters were estimated with a *parametric* method.
- We provide portfolio performance results for the global financial crisis year of 2008 using 14 different market-neutral trading algorithms and show that they perform significantly better in 2008 than they do in less volatile years between 2001 and 2007.

2 Selection Strategies

Constructing a profitable trading strategy always starts with the comprehensible sifting of investment options. In this paper, four types of quantitative selection techniques and their combinations will be discussed: minimum distance method (MDM), ADF test, two-way Granger causality test and the MFR method.

2.1 Minimum Distance Method

Gatev *et al.* (2006) test their pairs trading system over daily S&P500 data dating from 1962 to 2002. They select pairs of stocks using the MDM. The main idea to their selection criterion is to select pairs that have had similar historical price moves. According to law of one price theory (Coleman, 2009), similar securities would have similar prices. To start the process, it is assumed that all the prices are equal to 1.00 for the starting day. Then, a cumulative return index is generated for all stocks. To select pairs from this data set, the sum of squared deviations is used:

$$\gamma(X,Y) = \sum_{t=1}^{T} \left(C_X^t - C_Y^t\right)^2 \tag{1}$$

where C_X^t and C_Y^t are the cumulative return indices for assets X and Y at time t. The smaller value of $\gamma(X, Y)$ gives us the information that selected stocks have had similar price changes until time T.

2.2 ADF Test

In order to generate a profit in a pair-trade, the ratio of the prices, R_t, needs to have both a constant mean and a constant volatility over time. We use the well-known unit root test to check for weak stationarity, the existence of which proves that we have what we are looking for. For an autoregressive process AR(1) such as $\delta X_t = (\phi_1 - 1)X_{t-1} + \varepsilon_t$, and defining $a \equiv \phi_1 - 1$ the unit root test can be written as follows:

$$\mathrm{H}_0 : a = 0$$

$$\mathrm{H}_1 : a < 0$$

The term 'augmented' comes from the number of lagged values of the dependent variable (Dickey and Wayne, 1979). The number of lagged difference terms to include is determined empirically, the idea being to include enough terms so that the error term in the tested equation is serially uncorrelated. Tau statistics will be used to determine the passing pairs.

2.3 ADF Test Combined with Two-way Granger Causality

Our top concern is the risk that one takes when entering a pair-trade, which is the possibility of a structural breakdown of the mean-reverting-price-ratio property. This paper tries to minimize this risk by both choosing the pairs more carefully and using an *ad hoc* trade time-out date, which is the last trading day of the selected year.

This is how we choose the pairs. Note that the ADF test for R_t gives one of only two results, Pass or Fail. This means that we cannot sort the tau statistics as we did in the MDM. Because there were too many pairs that passed the ADF test, and because some of the selected pairs did perform poorly the year after they were selected, we decided that we needed additional testing.

This is where the Granger causality test in *both* directions comes in. As is well known, Granger causality does not mean causality in the logical sense.

Rather, 'P_A Granger-causes P_B' means the former can be used to predict the latter. Obviously, two-way Granger causality is stronger than one-way Granger causality. A pair selected as such makes it less likely for the aforementioned structural breakdown to take place before the trade is timed out at the end of the year.

2.4 Market Factor Ratio

As mentioned before, by opening one long and one short position (–A, B), any pairs trading strategy becomes a market-neutral strategy *to some extent*. That means not all pair-trades are 100 per cent market-neutral. Here, we investigate a method to measure the degree of market neutrality of the pair selected. This is done by picking pairs that have highly similar market exposures, or betas. The closer the betas are, the better the market-risk hedging is (Elton *et al.*, 2007). We generate an MFR criterion for each possible pair.

$$\mathrm{MFR} = \mathrm{abs}\left(\frac{\beta_1}{\beta_2}\right) - 1 \tag{2}$$

Then, the MFRs are listed in increasing order, and the top five market-neutral pairs are selected from this list. We choose the top five because we want to make a fair comparison with other pairs trading strategies, which also use a quota of five, the industry standard.

Note that this method aims at decreasing the market risk and thus tries to make profits based on mispricings of the elements of a pair. In other words, the *MFR method capitalizes on the hidden idiosyncratic risk of individual stocks that comprise the pair without worrying about where the market is headed.*

3 Trading Algorithms

3.1 Two-standard-deviation Rule

Let μ_t be the moving average and s_t be historical volatility (moving standard deviation) of the ratio R_t at time t. As expressed in Gatev *et al.* (2006), traders generally use a rule of thumb, namely the two-standard-deviation rule, in entering a pair-trade (–A, B) or (A, –B). The idea is to open the pair-trade when R_t increases (or diminishes) and hits the two-standard-deviation barrier $\mu_t + 2s_t$ (or $\mu_t - 2s_t$) and to close it when R_t returns to its moving average. We use the shorthand notation '2STD' for this specific trade rule. It is clear that this overly simplistic rule can be optimized by using more sophisticated quantitative tools. One such tool is to use the Vasicek stochastic differential equation model for R_t.

3.2 Implementation of Vasicek Model to Pairs Trading

In our market-neutral trading system, instead of using the moving average and historical volatility of R_t, we use θ and σ, the long-term mean and instantaneous standard deviation, respectively, that belong to the Vasicek model of R_t. Note that we estimate fresh values of θ and σ at every time step (every week) along the way.

Our research investigates a Vasicek model-based trading system that uses the two-standard-deviation rule to open a trade. (We use the shorthand notation 'V2STD' for this specific trade rule.) That is, we open the trade (–A, B) when $R_t \geq \theta + 2\sigma$. We close this trade when $R_t \leq \theta$. Likewise, we open the trade (A, –B) when $R_t \leq \theta - 2\sigma$. We close this trade when $R_t \geq \theta$.

Here, we shall elaborate the implementation of the Vasicek model to market-neutral trading. Mean reversion is a tendency for a stochastic process to remain near or tend to return over time to its long-run mean. As well-known examples, interest rates and implied volatilities can be given. In general, stock prices themselves do not tend to have mean reversion. The Vasicek model (Vasicek, 1977) is generally used for interest rate modeling, but it can easily be applied to other mean-reverting processes as well. This model assumes that a mean-reverting process has a stochastic differential equation in the form

$$\mathrm{d}R_t = \kappa(\theta - R_t) + \sigma \mathrm{d}W_t \tag{3}$$

where W_t is a Wiener process that models the continuous randomness of the system. θ is the long-term mean around which all future trajectories of R_t will evolve. κ is the speed of mean reversion. A very high κ can lead to fewer trading opportunities, whereas a very low one can lead to a more risky trading structure. σ is the instantaneous volatility, a very high value of which may easily lead to a risky trading system.

The evolutions of these parameters are of importance. When we solve the stochastic differential equation, we come to the result

$$R_t = R_0\mathrm{e}^{-\kappa t} + \theta(1 - \mathrm{e}^{1\kappa t}) + \sigma e^{-\kappa t}\int_0^t \mathrm{e}^{\kappa s}\mathrm{d}W_s \tag{4}$$

and the expected value or the mean is

$$E[R_t] = R_0\mathrm{e}^{-\kappa t} + \theta(1 - \mathrm{e}^{-\kappa t}) \tag{5}$$

with the variance

$$\frac{\sigma^2}{2\kappa}(1 - \mathrm{e}^{2\kappa}) \tag{6}$$

Thus, the long-term mean is

$$\lim_{x \to \infty} E[R_t] = \theta \tag{7}$$

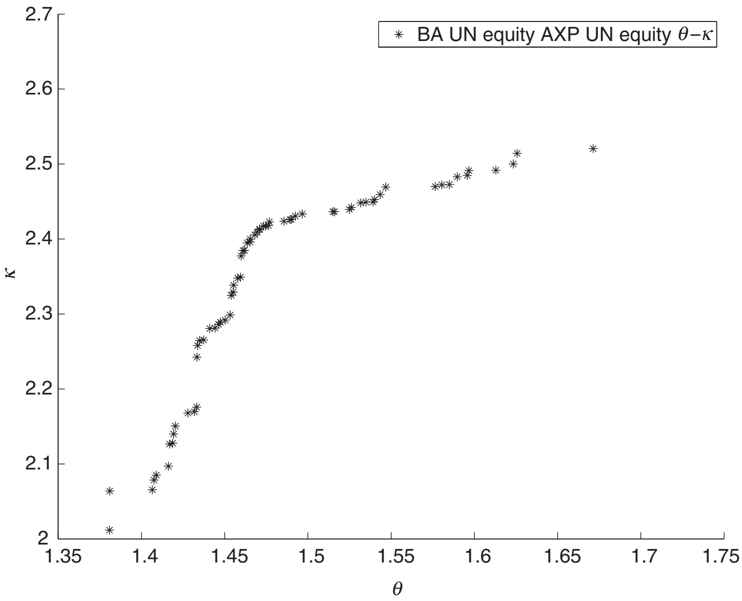

FIG. 1 κ vs. θ

and the long-term variance is

$$\frac{\sigma^2}{2\kappa} \tag{8}$$

In order to visualize how the Vasicek model works for the pair BA–AXP (Boeing and American Express) see Figs 1–8.

In Do *et al.* (2006), the authors use expectation maximization (Shumway and Stoffer, 1982) and the Kalman filter (Kalman, 1960) to estimate Vasicek parameters for a pairs trading system. However, this choice requires the need to make an assumption about the distribution of parameters. In our system, we do not make any such assumptions since we use the GMM, a non-parametric model.

4 GMM Estimation Technique

To explain the dynamic properties of econometric systems, parameter estimation procedures have crucial importance. The GMM was first introduced by Hansen (1982). GMM is a flexible tool used in a large number of econometric and economic models. By relying on gentle and convincing assumptions, GMM has had a significant impact on the theory and practice of econometrics. For the theory side, the main earning is that GMM provides a

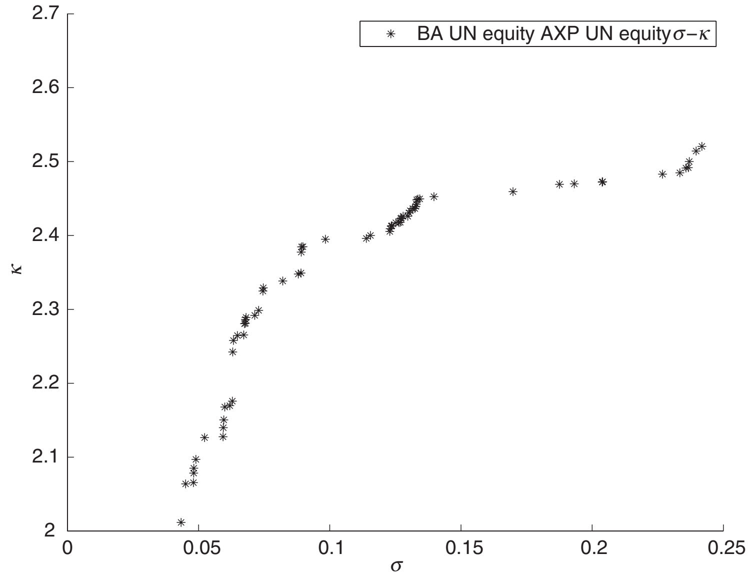

FIG. 2 κ VS. σ

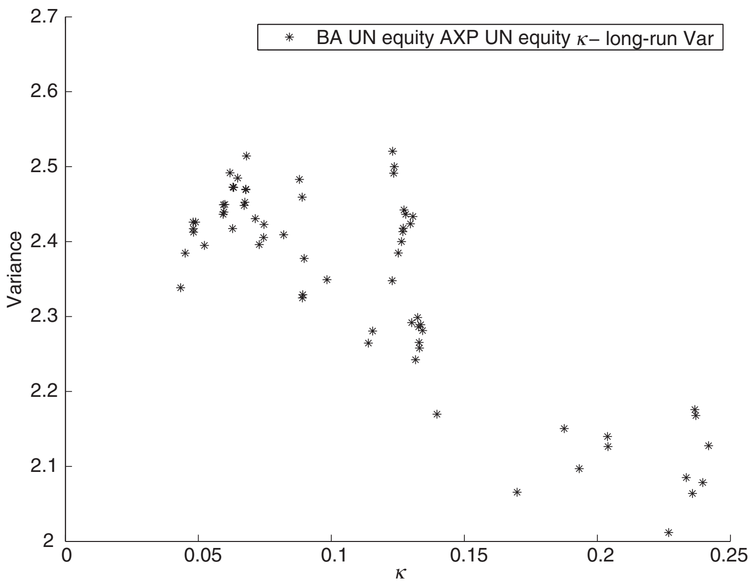

FIG. 3 Long-term Variance vs. κ

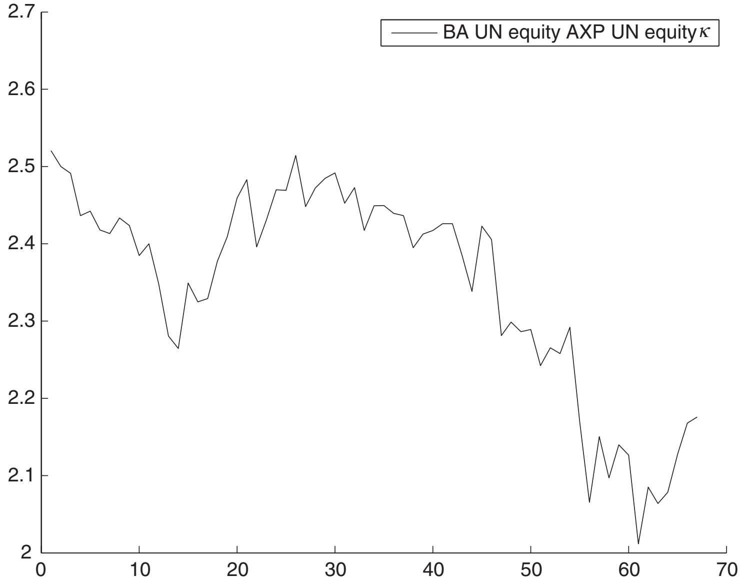

FIG. 4 κ vs. Time

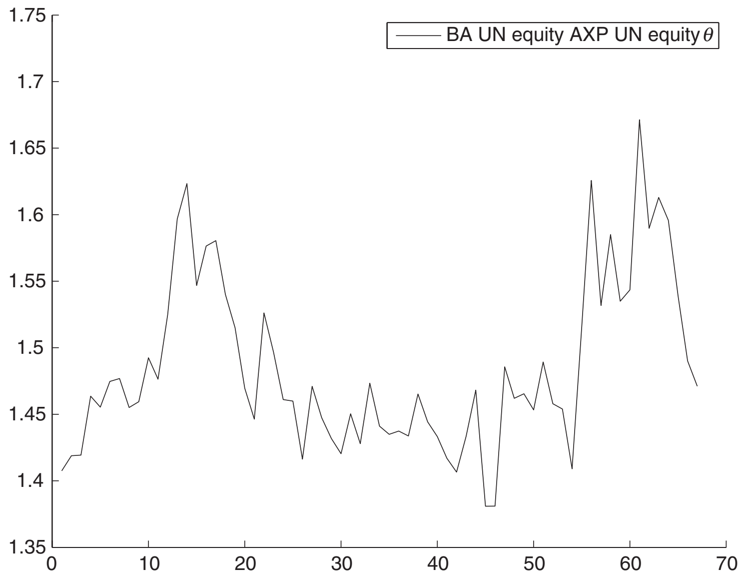

FIG. 5 θ vs. Time

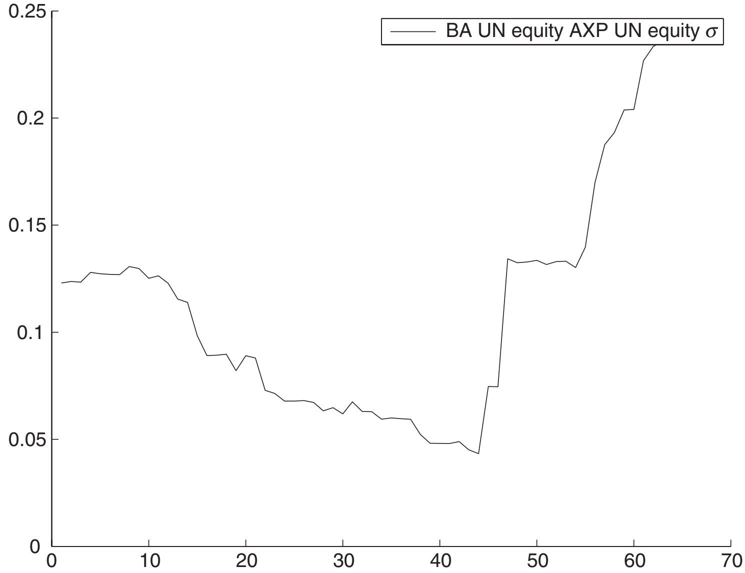

FIG. 6 σ vs. Time

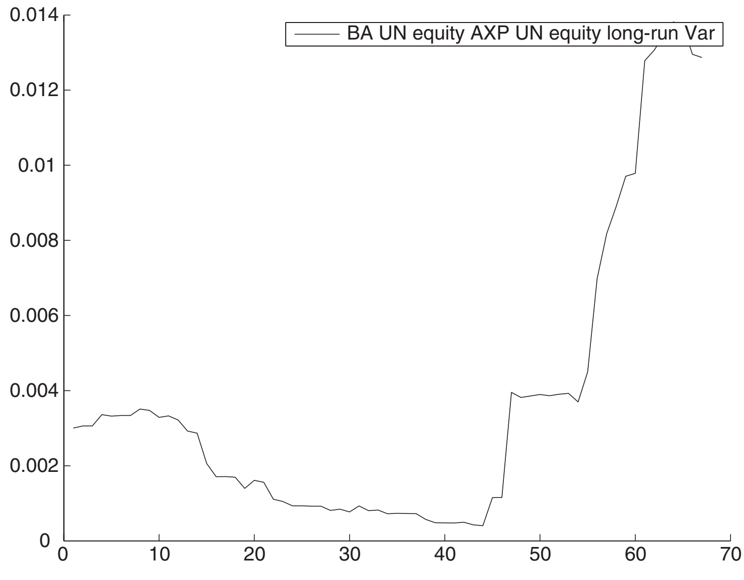

FIG. 7 Long-term Variance vs. Time

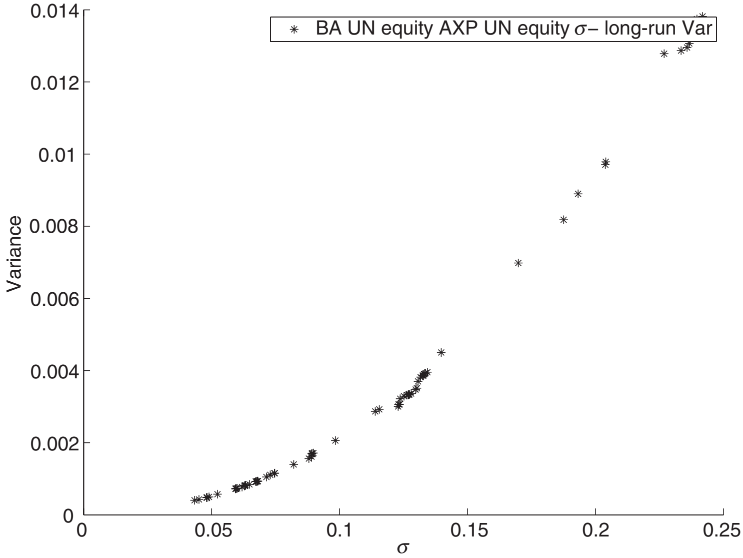

FIG. 8 Long-term Variance vs. σ

very general framework for considering issues of statistical consequence because it entails the solution to finding many estimators of interest in econometrics. For the practical side, unlike other methods such as the maximum likelihood process, it generates a computationally appropriate method of estimating non-linear dynamic models without making any assumptions on the probability distribution of the data. The only necessary input for GMM is the first few moments derived from the underlying model. This property makes GMM very useful in areas such as macroeconomics, finance, agricultural economics, environmental economics and labor economics.

To estimate the parameters of the Vasicek model explained earlier, we use GMM estimation. Our trading system will include the parameters θ and σ to make the trade decisions.

We assume modeling the ratio R_t with the Vasicek model and estimating the parameters θ and σ will give us dynamic information on the behavior of the pairs. (In order to compute the former two parameters, κ also needs to be estimated because it is one of the unknowns of the non-linear GMM system.)

To discretize the continuous Vasicek stochastic equation, the following steps are implemented (Vasicek, 1977). The continuous-time model is restated as

$$dR_t = \kappa(\theta - R_t) + \sigma dW_t \tag{9}$$

where R_t is the real data (the ratio of the selected pairs at a selected time t). So,

$$E[dR_t] = \kappa(\theta - R_t)dt \tag{10}$$

A discrete-time approximation is

$$R_t - R_{t-1} = \kappa(\theta - R_{t-1})[t-(t-1)] + \varepsilon_t \tag{11}$$

Let $Y_t = R_t - R_{t-1}$ and $S = -\kappa$ and $Q = -S\theta$. Thus,

$$Y_t = \frac{1}{52}(Q + SR_{t-1}) + \varepsilon_t \tag{12}$$

Note that we use 1/52 for $t - (t - 1)$ because we use weekly data. So,

$$e_1(t) = \frac{1}{52}(Q + SR_{t-1}) \tag{13}$$

$\mathrm{Var}[\mathrm{d}R_t] = \sigma^2\mathrm{d}t$. Because $E[\mathrm{d}W_t^2] = \mathrm{d}t$

$$e_2(t) = e_1^2(t) - \frac{\sigma^2}{52} \tag{14}$$

and g is defined as

$$g = [e_1, e_2] \tag{15}$$

Now, the goal is to estimate the unknown parameters, Q, S and σ, by minimizing the quadratic form $g^{\mathrm{T}}Wg$, where W is a weight matrix that considers the variances of the moments and gives more positive weighting to the component of g that has a smaller variance. The optimization is done iteratively, using the fmincon function of MATLAB. This function is an efficient optimizer for non-linear systems with constraints.

5 Application Methodology

This section describes the methodology used for the analysis in this paper. First, it introduces the disjoint training and testing periods used in the experiments, and then it introduces the algorithms used for pair selection and trading.

5.1 Training and Testing Periods

We first define two consecutive time periods as training and testing. The training period is a preselected period where the parameters of the experiment are calculated and frozen. Immediately after the training period, the testing period follows, where we run the experiments with these frozen parameters. Note that pairs are also treated as parameters in our trading system. We use one year for training and the consequent year for testing.

In our analysis, we first select pairs and then make trading decisions using one-step-ahead (one-week-ahead) estimates of the parameters of the underlying Vasicek model. To generate a one-step-ahead forecast, we need to specify a fixed moving window length similar to a window length used in calculating moving averages.

Consequently, our training period needs to find answers to two questions.

- What are the best pairs for trading?
- What is the optimum window length?

We first select pairs with a selection algorithm and then calibrate the optimum window length. That is, we scan the same training period twice, once for pair selection and once for window length optimization.

We select pairs by a combination of the following methods mentioned earlier: MDM, MFR, ADF test, and the two-way Granger causality test, abbreviated as G. We use a plus sign for a combination of two methods.

To be specific, we select our pairs using seven different methods: {MFR}, {ADF + G + MFR}, {ADF + MFR}, {G + MFR}, {G}, {ADF + G} and {MDM}. Note that for all selection methods that involve {MDM} or {MFR}, we select the top five pairs from a sorted list of minimum distance or minimum market factor ratio, respectively, and create an equally weighted portfolio in each case. For selection methods involving {G} but not {MFR}, the top five passing pairs are selected, where the sort is based on the sum of p values of the two-way Granger causality tests.

Window length optimization is actually an optimization based on profit. With each selected pair, we trade with 24, 36, 48, 60 and 72 weeks of window length in the training period. The window length with the highest cumulative profit is selected as the optimum window length.

We also have a trade time-out date, which is 30 December of each year. Once the trade time-out date is reached, the pair-trade is closed no matter what the profit or loss is. The industry has similar time-out mechanisms. One would close a pair-trade if a certain amount of time has passed or a certain fixed date is reached.

Note that we do not use any stop-loss or take-profit parameters, which is not the usual way the industry does pairs trading. Most of the time, the industry uses *ad hoc* parameters for stop-loss and take-profit.

6 Experiments and Analysis

6.1 Data and Coding Infrastructure

We have downloaded end-of-week price data for stocks that comprise the Dow Jones 30 index[1] from Thompson Reuters Datastream (http://online.thomsonreuters.com/datastream/). In all our experiments, the weeks between the last trading day of the 26th week of the year N and the last trade date of the year $(N+1)$ are selected as the training period, whereas the weeks between 2 January $(N+2)$ and 30 December $(N+2)$ comprise the testing period, where

[1]Note that, after this data set was decided upon, two members of Dow Jones 30 were replaced by others. On 8 June 2009, GM and Citigroup were replaced by The Travelers Companies and Cisco Systems, respectively.

$N = 1999, \ldots, 2006$. As mentioned before, we use weekly data and select pairs from stocks that comprise the Dow Jones 30 index. Consequently, we do the analysis for $\binom{30}{2} = 435$ possible pairs.

6.2 *Results from Each of the Selection Methods*

For our improvement on the ADF test, {ADF + G}, we record that 15 per cent of the possible total pairs passed through ADF test, but only 12 per cent of them achieved successful results on the two-way Granger causality test (although not all managed to make the quota of five). Thus, the percentage of the pairs that passed both the ADF and the two-way Granger causality was only 1.8 per cent. Also, note that in 88 per cent of the pairs selected by the ADF test, the stationarity of the price ratio R_t was reinforced by only *one* component of the pair, whereas in the remaining 12 per cent it is reinforced by both members.

The pairs that were selected by each pair selection method are given in Tables 1–7. The number of distinct pairs selected by seven selection methods are given in Table 8. Also, the industries of the selected pairs are listed in the Appendix.

TABLE 1
ADF TEST–GRANGER CAUSALITY SELECTION RESULTS

Year	*Pair 1*	*Pair 2*	*Pair 3*	*Pair 4*	*Pair 5*
2001	JPM–AA	JNJ–C	DIS–DD	CVX–BAC	XOM–AXP
2002	MSFT–MMM	MSFT–DD	MSFT–C	MSFT–AA	T–KO
2003	IBM–BA	CVX–BA	UTX–DD	INTC–HD	GE–BA
2004	UTX–BA	PG–IBM	MSFT–BAC	PFE–DD	T–BAC
2005	JNJ–BAC	JPM–AA	MCD–DIS	MCD–DD	DIS–AXP
2006	VZ–DIS	IBM–CVX	PG–AXP	JPM–BA	XOM–IBM
2007	INTC–GE	UTX–JNJ	INTC–IBM	JNJ–DIS	JNJ–INTC
2008	MRK–MCD	T–GE	PFE–AA	GE–CAT	BA–AXP

TABLE 2
MFR SELECTION RESULTS

Year	*Pair 1*	*Pair 2*	*Pair 3*	*Pair 4*	*Pair 5*
2001	MMM–JPM	XOM–CVX	T–HPQ	HPQ–AXP	PG–MSFT
2002	CVX–C	T–PFE	HD–BA	T–GE	XOM–PFE
2003	UTX–AXP	VZ–C	MSFT–JPM	MSFT–AA	T–GE
2004	BAC–AXP	XOM–CVX	VZ–JPM	GE–AA	JPM–CAT
2005	WMT–PG	CVX–BA	DD–BAC	JPM–HD	MRK–JPM
2006	MSFT–HPQ	XOM–CVX	KO–DD	INTC–DIS	PFE–DIS
2007	KO–JPM	UTX–JNJ	CVX–CAT	DIS–AA	XOM–CAT
2008	WMT–JPM	HPQ–DD	XOM–CVX	HD–DIS	UTX–CAT

Table 3
ADF Test—Sorted by MFR Selection Results

Year	*Pair 1*	*Pair 2*	*Pair 3*	*Pair 4*	*Pair 5*
2001	JNJ–C	HPQ–HD	DD–C	T–C	JPM–HD
2002	HD–CVX	HD–C	PFE–DD	XOM–DD	DD–CVX
2003	UTX–AXP	VZ–C	CVX–BA	WMT–JNJ	PFE–C
2004	HPQ–DIS	JPM–HD	HD–BA	DD–BAC	UTX–BA
2005	UTX–AXP	DD–AXP	UTX–BAC	MCD–DIS	WMT–C
2006	C–BAC	XOM–PG	UTX–PG	WMT–BAC	UTX–AXP
2007	KO–JPM	UTX–JNJ	VZ–GE	MCD–HPQ	MMM–BA
2008	XOM–CVX	DD–BAC	MRK–MCD	GE–AA	JPM–DD

Table 4
Granger Causality Test—Sorted by MFR Selection Results

Year	*Pair 1*	*Pair 2*	*Pair 3*	*Pair 4*	*Pair 5*
2001	JNJ–C	HPQ–HD	XOM–PFE	JPM–DD	T–JNJ
2002	HD–C	MCD–BAC	T–CVX	VZ–CVX	T–KO
2003	CVX–BA	JPM–AA	JPM–GE	JPM–CAT	JPM–AXP
2004	CVX–BAC	VZ–CVX	UTX–BA	BA–AA	PFE–DD
2005	VZ–AA	MCD–DIS	WMT–C	XOM–JNJ	JPM–AA
2006	XOM–CVX	UTX–PG	UTX–AXP	KO–AXP	PG–AXP
2007	UTX–JNJ	XOM–JNJ	MMM–IBM	MSFT–AA	JPM–BAC
2008	MRK–MCD	T–GE	C–AA	JPM–AA	C–AXP

Table 5
Granger Causality Test Selection Results

Year	*Pair 1*	*Pair 2*	*Pair 3*	*Pair 4*	*Pair 5*
2001	JPM–AA	JNJ–C	DIS–DD	CVX–BAC	MMM–DD
2002	MSFT–MMM	WMT–HD	WMT–UTX	UTX–MMM	MSFT–DD
2003	JPM–CAT	JPM–C	MRK–JPM	VZ–JPM	JPM–AA
2004	UTX–BA	MSFT–CVX	VZ–CVX	PG–IBM	BA–AA
2005	JNJ–BAC	JPM–AA	MCD–DIS	MCD–DD	DIS–AXP
2006	JPM–AA	XOM–CVX	VZ–DIS	WMT–AA	IBM–CVX
2007	INTC–GE	UTX–JNJ	IBM–HD	MMM–IBM	INTC–IBM
2008	MSFT–KO	MRK–MCD	C–AA	CAT–C	JPM–AA

7 Results of the Trading Methods

After selecting the pairs in the training period, we run a profit-based window length optimization for each trading system as we discussed earlier. Then with this optimum window length, we start trading with each of these portfolios. We use two different trading algorithms 2STD and V2STD. The cumulative profit of each portfolio for each algorithm is shown in Figs 9 and 10.

Tables 9–13 show the annual net returns (net in the sense that *only* the average annual US risk-free rate (http://www.federalreserve.gov/fomc/

TABLE 6
ADF TEST–GRANGER CAUSALITY—SORTED BY MFR SELECTION RESULTS

Year	*Pair 1*	*Pair 2*	*Pair 3*	*Pair 4*	*Pair 5*
2001	JNJ–C	HPQ–HD	JPM–DD	HD–DD	MCD–AA
2002	HD–C	MCD–BAC	T–CVX	VZ–CVX	T–KO
2003	CVX–BA	GE–BA	UTX–DD	INTC–HD	IBM–BA
2004	UTX–BA	PFE–DD	MSFT–BAC	T–BAC	PG–IBM
2005	MCD–DIS	WMT–C	XOM–JNJ	JPM–AA	JNJ–BAC
2006	UTX–PG	UTX–AXP	PG–AXP	VZ–DIS	VZ–KO
2007	UTX–JNJ	JPM–BAC	WMT–UTX	DIS–DD	INTC–DIS
2008	MRK–MCD	T–GE	PFE–AA	GE–CAT	T–MMM

TABLE 7
MDM SELECTION (GATEV *ET AL.*, 2006) RESULTS

Year	*Pair 1*	*Pair 2*	*Pair 3*	*Pair 4*	*Pair 5*
2001	DD–AA	CAT–AA	GE–AXP	KO–JNJ	WMT–CAT
2002	MMM–DD	T–MCD	XOM–DD	XOM–MMM	PFE–DD
2003	UTX–DD	VZ–C	XOM–WMT	MMM–BAC	XOM–CAT
2004	XOM–PG	C–AXP	PG–IBM	DIS–C	PG–CVX
2005	XOM–CVX	JNJ–BAC	VZ–BAC	GE–BAC	PFE–KO
2006	GE–BAC	C–BAC	GE–C	UTX–PG	T–HD
2007	GE–C	IBM–DD	IBM–AXP	UTX–JPM	PG–JNJ
2008	DIS–DD	MMM–IBM	XOM–UTX	XOM–HPQ	JPM–DD

TABLE 8
THE NUMBER OF DISTINCT PAIRS SELECTED BY SEVEN SELECTION METHODS

	Six methods	*MDM*	*Overlap*
2001	20	5	1
2002	21	5	2
2003	19	5	2
2004	18	5	1
2005	16	5	1
2006	19	5	2
2007	20	5	0
2008	19	5	1

fundsrate.htm) is deducted) and the Sharpe ratios for each trading market-neutral trading system.

8 CONCLUSIONS

Market-neutral trading strategies exploit market inefficiencies, or mispricings in a pair of similar stocks, which are more commonplace in a global crisis allowing more trading possibilities to emerge at bad times. Moreover, there

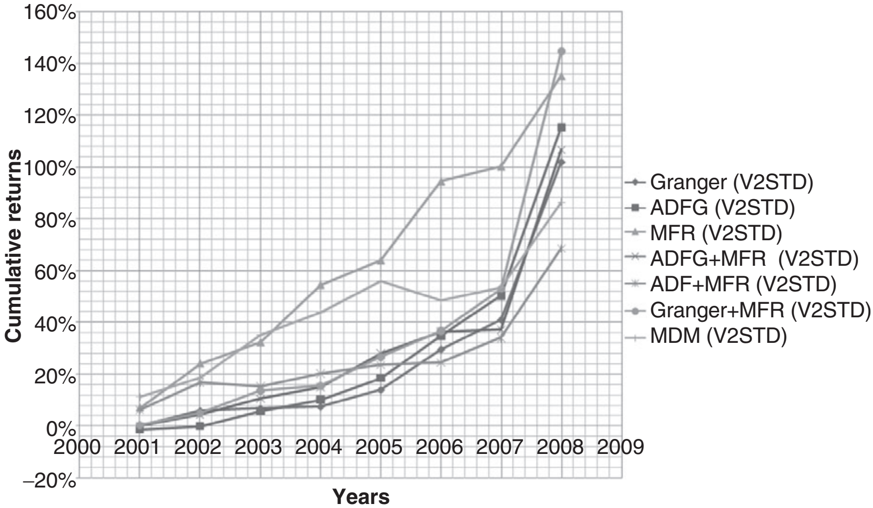

FIG. 9 Cumulative Return Graphs of V2STD Trading Methods

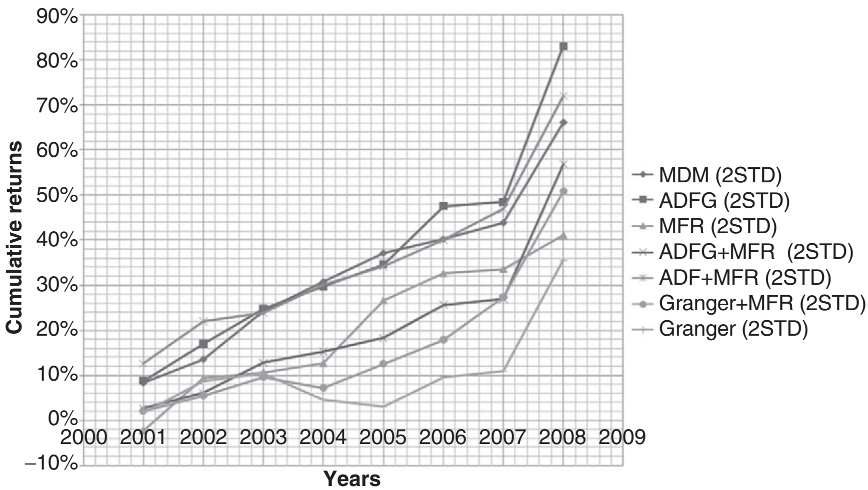

FIG. 10 Cumulative Return Graphs of 2STD Trading Methods

are fewer market participants, which reduces competition. Therefore it is not surprising that market-neutral trading performs best during the most severe market conditions. In this paper, we have shown empirical proof that supports the above statement.

There were three other important conclusions to be drawn from our experiments.

TABLE 9
AVERAGE ANNUAL US RISK-FREE RATES (http://www.federalreserve.gov/fomc/fundsrate.htm)

Year	*2001*	*2002*	*2003*	*2004*	*2005*	*2006*	*2007*	*2008*
r_f	3.91	1.67	1.12	1.34	3.19	4.96	5.05	2.10

TABLE 10
NET RETURNS AND SHARPE RATIOS

	1	*2*	*3*	*4*
Year	*MDM (2STD)*	*MDM (V2STD)*	*ADF + G (2STD)*	*ADF + G (V2STD)*
2001	4.38	7.15	4.83	–5.30
2002	3.19	4.95	5.86	–0.38
2003	8.25	12.94	5.53	4.75
2004	3.93	5.13	2.63	2.73
2005	1.70	5.12	0.66	4.51
2006	–2.71	–9.68	4.61	8.71
2007	–2.45	–1.63	–4.45	6.53
2008	13.40	19.29	21.26	41.14
STD	5.32	8.70	7.36	14.13
Average	3.71	5.41	5.12	7.84
Sharpe	0.70	0.62	0.70	0.55

TABLE 11
NET RETURNS AND *S*HARPE RATIOS

	5	*6*	*7*	*8*
Year	*MFR (2STD)*	*MFR (V2STD)*	*ADF + G + MFR (V2STD)*	*ADF +* G *+ MFR (2STD)*
2001	–6.20	3.09	–3.97	–1.21
2002	10.27	14.25	2.72	1.66
2003	0.07	5.55	4.75	5.19
2004	0.42	15.37	2.73	0.87
2005	9.20	2.93	8.14	–0.52
2006	–0.23	13.72	1.60	1.20
2007	–4.29	–2.10	–4.48	–4.03
2008	3.52	15.35	48.61	21.43
STD	5.85	6.93	17.13	7.87
Average	1.60	8.52	7.51	3.07
Sharpe	0.27	1.23	0.44	0.39

1. For our improvement on the ADF test we have observed that 15 per cent of the possible total pairs passed the ADF test, but only 12 per cent of them achieved successful results on the two-way Granger causality test (though not all managed to make the quota of five). Thus, the percentage of the pairs that passed both the ADF and the two-way Granger causality

Table 12
Net Returns and Sharpe Ratios

Year	9 ADF + MFR (2STD)	10 ADF + MFR (V2STD)	11 G + MFR (2STD)	12 G + MFR (V2STD)
2001	8.68	2.19	−1.84	−3.65
2002	6.66	8.37	1.66	2.94
2003	0.40	−2.42	2.86	7.20
2004	3.62	2.97	−3.59	0.23
2005	0.06	−0.42	1.83	6.54
2006	−0.58	−4.24	−0.29	2.91
2007	−0.17	2.55	3.00	6.83
2008	14.96	23.72	16.31	57.97
STD	5.55	8.81	6.05	19.69
Average	4.20	4.09	2.49	10.12
Sharpe	0.76	0.46	0.41	0.51

Table 13
Net Returns and Sharpe Ratios

Year	13 G (2STD)	14 G (V2STD)
2001	−2.20	−3.83
2002	5.28	4.10
2003	0.39	−0.25
2004	−6.66	−0.70
2005	−4.64	2.86
2006	1.47	8.74
2007	−3.85	3.80
2008	20.06	41.09
STD	8.50	14.29
Average	1.23	6.98
Sharpe	0.14	0.49

was only 1.8 per cent. In other words, in 88 per cent of the pairs selected by the ADF test, the stationarity of the price ratio R_t was reinforced by only one component of the pair, whereas in the remaining 12 per cent it is reinforced by both members.

2. The V2STD trading rule performs better than the simple 2STD trading rule when the performance criterion is the average return over eight years. This statement was not true when the performance criterion was the Sharpe ratio over eight years (in which we used annual returns).
3. In 2008, the first year of the global financial crisis, the pairs that were selected using some combination of the two-way Granger causality rule

and traded with the V2STD rule outperformed all competing models considered in this paper, where the performance criterion was only the annual return. Note that more than 40 per cent returns were observed in each and every one of these cases.

Appendix

Table A1
Members of the Dow Jones 30 Index

Symbol	*Industry*	*Company*
MMM	Conglomerate	3M
AA	Aluminum	Alcoa
AXP	Consumer finance	American Express
T	Telecommunication	AT&T
BAC	Banking	Bank of America
BA	Aerospace and defense	Boeing
CAT	Construction and mining equipment	Caterpillar
CVX	Oil and gas	Chevron Corporation
C	Financial services	Citygroup
KO	Beverages	Coca-Cola
DD	Chemical industry	DuPont
XOM	Oil and gas	ExxonMobil
GE	Conglomerate	General Electric
HPQ	Technology	Hewlett-Packard
HD	Home improvement retailer	The Home Depot
INTC	Semiconductors	Intel
IBM	Computers and technology	IBM
JNJ	Pharmaceuticals	Johnson & Johnson
JPM	Banking	JPMorgan Chase
KFT	Food processing	Kraft Foods
MCD	Fast food	McDonald's
MRK	Pharmaceuticals	Merck
MSFT	Software	Microsoft
PFE	Pharmaceuticals	Pfizer
PG	Consumer goods	Procter & Gamble
GM	Automotive	General Motors
UTX	Conglomerate	United Technologies Corporation
VZ	Telecommunication	Verizon Communications
WMT	Retail	Wal-Mart
DIS	Broadcasting and entertainment	Walt Disney

References

Avellaneda, M. and Lee, J.-H. (2008). 'Statistical Arbitrage in the U.S. Equities Market' (11 July); available at SSRN: http://ssrn.com/abstract=1153505.

Beliossi, G. (2002). 'Market Neutral Strategies', *Journal of Alternative Investments*, Vol. 5, No. 2, pp. 1–4.

Coleman, A. (2009). 'Storage, Slow Transport, and the Law of One Price: Theory with Evidence from Nineteenth-century U.S. Corn Markets', *Review of Economics and Statistics*, Vol. 91, No. 2, pp. 332–350.

Dickey, D. A. and Wayne, A. F. (1979). 'Distribution of the Estimators for Autoregressive Time Series with a Unit Root', *Journal of the American Statistical Association*, Vol. 74, pp. 427–431.

Do, B., Faff, R. and Hamza, K. (2006). 'A New Approach to Modeling and Estimation for Pairs Trading', *Working Paper*, Monash University, 29 May.

Elliott, R. J., van der Hoek, J. and Malcolm, W. P. (2005). 'Pairs Trading', *Quantitative Finance*, Vol. 5, No. 3, pp. 271–276.

Elton, E. J., Gruber, M. J., Brown, S. J. and Goetzman, W. N. (2007). *Modern Portfolio Theory and Investment Analysis*, 7th edn, New York, Wiley.

Engle, R. F. and Granger, C. W. J. (1987). 'Cointegration and Error Correction: Representation, Estimation, and Testing', *Econometrica*, Vol. 55, pp. 251–276.

Gatev, E., Goetzmann, W. N. and Rouwenhorst, K. G. (2006). 'Pairs Trading: Performance of a Relative-value Arbitrage Rule', *Review of Financial Studies*, Vol. 19, No. 3, pp. 797–827.

Granger, C. W. J. (1969). 'Investigating Causal Relations by Econometric Models and Cross-spectral Methods', *Econometrica*, Vol. 37, No. 1, pp. 424–438.

Hansen, L. P. (1982). 'Large Sample Properties of Generalized Method of Moments Estimators', *Econometrica*, Vol. 50, No. 4, pp. 1029–1054.

Herlemont, D. (2003). 'Pairs Trading, Convergence Trading, Cointegration', *Working Paper*.

Jegadeesh, N. and Titman, S. (1993). 'Returns to Buying Winners and Selling Losers: Implications for Stock Market Efficiency', *Journal of Finance*, Vol. 48, No. 1, pp. 65–91.

Kalman, R. E. (1960). 'A New Approach to Linear Filtering and Prediction Problems', *Journal of Basic Engineering*, Vol. 82, No. 1, pp. 35–45.

Kovajecz, K. A. and Odders-White, E. R. (2004). 'Technical Analysis and Liquidity Provision', *Review of Financial Studies*, Vol. 17, pp. 1043–1107.

Lehmann, B. (1990). 'Fads, Martingales, and Market Efficiency', *Quarterly Journal of Economics*, Vol. 105, No. 1, pp. 1–28.

Perlin, M. (2007). 'Evaluation of Pairs Trading Strategy at the Brazilian Financial Market' (November); available at SSRN: http://ssrn.com/abstract=952242.

Shumway, R. and Stoffer, D. (1982). 'An Approach to Time Series Smoothing and Forecasting Using the EM Algorithm', *Journal of Time Series Analysis*, Vol. 3, No. 4, pp. 253–264.

Vasicek, O. (1977). 'An Equilibrium Characterization of the Term Structure', *Journal of Financial Economics*, Vol. 5, No. 1, pp. 177–188.

Vidyamurthy, G. (2004). *Pairs Trading: Quantitative Methods and Analysis*, 1st edn, New York, Wiley.